

Modified Grey Wolf Optimizer with Sparse Autoencoder for Financial Crisis Prediction in Small Marginal Firms

¹Dr. Rajib Bhattacharya

¹Associate Professor, Department of PGDM, Globosyn Business School, Mouza Chandi, PS Bishnupur JL No. 101, Kolkata, West Bengal, India.
rajib.learner@gmail.com

²Kafila

²Assistant Professor, School of Business, SR University, Warangal, Telangana, India.
Kafila@sru.edu.in

³Dr Somanchi Hari Krishna

³Associate Professor, Department of Business Management, Vignana Bharathi Institute of Technology, Aushapur Village Ghatkesar Mandal, Malkangiri Medichal, Telangana, India,
harikrishnasomanchi@gmail.com

⁴Dr Bhadrappa Haralayya

⁴Professor and HOD, Department of MBA, Lingaraj Appa Engineering College Bidar, Karnataka, India -585403,
bhadrappabhavimani@gmail.com

⁵Dr. Pooja Nagpal

⁵Associate Professor, Department of Management Studies, International School of Business Excellence ISME, Bangalore, Karnataka, India.
pooja.nagpal@gmail.com, Orcid id: 0000-0003-0772-0943

⁶Chitsimran

⁶Research Scholar, Mittal School of Business, Lovely Professional University, Phagwara, Punjab, India.
chitsimran24@gmail.com

Abstract— Small marginal firms play an important role in the economy, and their failure has widespread consequences. As a result, it is critical that small marginal firms be capable of predicting financial crises in order to mitigate their negative impact. Financial Crisis Prediction (FCP) is the process of recognizing the possibility of a future financial crisis. FCP is an important task for financial institutions, policymakers, and investors because it helps them to prepare for and mitigate the negative impact of financial crisis. Machine learning (ML) approaches are used to forecast financial crisis in small and marginal firms. It is completed by training a model on historical data and using it to predict the likelihood of a future financial crisis. As a result, this article proposes a Modified Grey Wolf Optimizer with Sparse Autoencoder (MGWO-SAE) for predicting financial crisis situations in small marginal firms. The goal of the MGWO-SAE technique is to forecast financial crisis effectively in small marginal firms. To accomplish this, the proposed MGWO-SAE technique employs data preprocessing to convert the input financial data into the correct format. The MGWO-SAE technique employs the SAE classification method for prediction. The MGWO algorithm is used for hyperparameter tuning model to improve the performance of the SAE technique. The MGWO-SAE technique's experimental result analysis is tested on a financial dataset. The experimental results demonstrated the advantages of the MGWO-SAE technique over other models.

Keywords— Small marginal firms; Financial crisis; Predictive models; Machine learning; Grey wolf optimizer

I. INTRODUCTION

The financial crisis means a significant deterioration in indicators like asset prices, interest rates, solvency, and exchange rates [1]. During the catastrophic crisis, lot of

companies become insolvent and stock prices fall abruptly, which is a huge loss to all participants in the financial market. With the rapid progression of financial incorporation, the financial security of many nations has influenced the speed of growing economy of other nations [2]. This attention was mainly directed to the concern that financial crisis excessively affects small and medium-sized enterprises (SMEs which can be considered an important source of employment, entrepreneurship, economic growth, and innovation [3]. At the time of the crisis, private SMEs encounters higher costs and additional restrictions in accruing access to finance relative to large, listed firms [4]. Such complexities arise from the different features of private firms; for instance, difficulty to access the public capital markets, higher failure rates, shorter track records, greater data asymmetry, less opportunities were to owners-managers for wealth diversification, and commonly lower collateral availability [5].

The Machine Learning (ML) techniques allow to manage numerous known difficulties in the literature on macroeconomic early warning mechanisms [6]: Initially, the dynamics preceding fiscal crises were extremely complex, and simple threshold or linear methods, while offering an intuitive narrative, may have a hard time capturing like complexities adequately [7]. Then, provided that the small sample sizes in macroeconomic panels, with at best a few thousand (correlated) observations, it is simple to find paradigms—narratives to explain past crises—that are spurious and will not be appropriate for forecasting future crises. This latter concern [8], the risk of overfitting (of explaining paradigms that are specific to estimation sample and do not generalize to other samples), is predominantly pressing for the techniques most popular between applied economists, maximal - least-squares

likelihood and which, by definition, try to fit the data they are “tuned to generate unbiased estimates of coefficients than reducing forecasting error” [9]. The two concerns regarding overfitting and underfitting constitute a renowned trade-off for any prediction method: forecasting fiscal crises need complicated methods; but by including complexity we might raise the risk of overfitting. ML techniques try to optimally solve the trade-off between overfitting and underfitting, which has made them method of choice for prediction problems [10].

This article designs a Modified Grey Wolf Optimizer with Sparse Autoencoder (MGWO-SAE) for Financial Crisis Prediction in Small Marginal Firms. The presented MGWO-SAE technique uses data preprocessing to transform the input financial data into a proper format. For prediction purposes, the MGWO-SAE technique uses SAE classification method. To enhance the performance of the SAE model, the MGWO algorithm is employed for hyperparameter tuning model. The experimental result analysis of the MGWO-SAE technique is tested on financial dataset.

II. RELATED WORKS

Zhang et al. [11] present a k-fold RF system integrates with time series analysis method as primary warning system for corporation FCP. This method proceeds benefit from the capability of time series analysis method for making short-term predictions of previous information and utilizes the time series analysis method for making predictions of recently created financial index data. Chen and Wu [12] exploit the textual financial data—letters to shareholders to present a structure for corporation FCP before standard typical numerical financial ratios. According to the financial crisis feature term library, queen GA and SVM are utilized for classifying letters to shareholders.

Batuman et al. [13] inspect the influence of global financial crisis on causes of corporate cash holdings and adjustments nearby target cash level utilizing an instance of Eastern European firms. Furthermore, the authors determine significantly low alteration speed for attaining the optimum cash level from the post-crisis epoch. In [14], the authors executed the forward-looking conditional value at risk (COVaR) as market-based systemic risk measure for analyzing the systemic threat from the developing market. Due to the exposure network, 2 novel network-based indices can be established. In addition, assuming the necessity of novel indices on values of forward-looking CoVaR, it can be executed as the indicator for the presented primary warning scheme for predicting downturns and crises from the firm level and entire markets.

Tyagi and Boyang [15] examine an intelligent IoT-assisted FCP method by utilizing meta-heuristic techniques. Initially, the financial data of enterprises can be collected by utilizing IoT devices such as laptops, smartphones, and so on. Afterward, the quantum artificial butterfly optimized (QABO) system for FS was executed for selecting a better feature set. Next, the LSTM with RNN methods is utilized for classifying the gathered financial data. In [16], for selecting suitable features, the authors present a hybrid filter-wrapper system. During the filter stage, the authors utilize the Spearman correlation and information gain are estimation conditions. During the wrapper stage, the authors combine the water wave

optimizer (WWO) that was modified based on the solution space of FS model.

III. THE PROPOSED MODEL

In this article, we have developed a new MGWO-SAE model for automated FCP in small marginal firms. The goal of the MGWO-SAE technique lies in the effectual forecasting of financial crisis in the small marginal firm. Fig. 1 represents the overall flow of MGWO-SAE approach.

A. Data Preprocessing

The presented MGWO-SAE technique uses data preprocessing to transform the input financial data into a proper format. This min–max normalized system is to scale every element feature value from the scope of [0, 1] based on the subsequent min–max normalized formula.

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

In which x stands for the source and x' denotes the normalized values correspondingly.

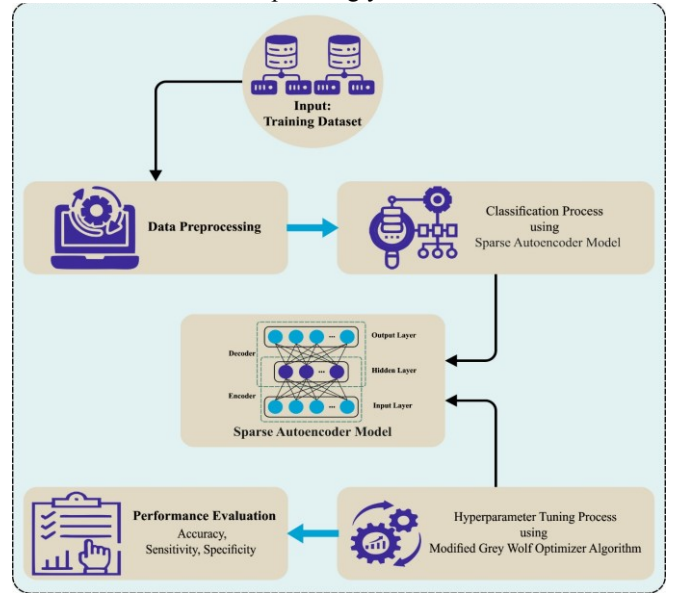


Fig. 1. Overall flow of MGWO-SAE approach

B. Prediction using SAE Model

For prediction purposes, the MGWO-SAE technique uses SAE classification model. The SAE with smoothed l_1 regularization method gives a labelled m amount of trained set records are $\{(x_l^{(1)}, y^{(1)}), (x_l^{(2)}, y^{(2)}), \dots, (x_l^{(m)}, y^{(m)})\}$ [17]. Where input feature vector represents $x_l^{(i)} \in \mathbb{R}^n$ (“ l ” specifies labelled records), $y^i \in \{+1, -1\}$ shows the dual classification label or $y^i \in \{1, 2, 3, \dots, C\}$ indicates multiple classification label. Consider that $x_u^{(1)}, x_u^{(2)}, x_u^{(3)}, \dots, x_u^{(m)} \in \mathbb{R}^n$ denotes the unlabelled sample produced by separating the label from the training set. To attain a suitable depiction and low dimensionality of input trained subset, $x_l^{(1)}, x_l^{(2)}, x_l^{(3)}, \dots, x_l^{(m)} \in \mathbb{R}^n$, we feed unlabelled trained instances $x_u^{(1)}, x_u^{(2)}, x_u^{(3)}, \dots, x_u^{(m)} \in \mathbb{R}^n$ in to present SAE technique. It is used for recreating and learning the training input data.

Afterward learning the optimum value of W and b using SAE on unlabelled dataset x_u . It learns and reconstructs the output value $(\hat{x}_u^{(1)}, \hat{x}_u^{(2)}, \hat{x}_u^{(3)}, \dots, \hat{x}_u^{(m)})$ to be equivalent to its input sample $(x_u^{(1)}, x_u^{(2)}, x_u^{(3)}, \dots, x_u^{(m)})$. Then, get a new and lower dimensional demonstration $\{(h_u^{(1)}, y^{(1)}), (h_u^{(2)}, y^{(2)}), (h_u^{(3)}, y^{(3)}), \dots, (h_u^{(m)}, y^{(m)})\}$. Where the original input sample is restored by respective sample of activation h . The procedure of reduction of dimensionality and feature extraction contains two stages. The encoding procedure map the input dataset x_i into the hidden unit demonstration as follows:

$$Z = h(WX + b) \quad (2)$$

$$\hat{X} = g(W'Z + b') \quad (3)$$

In the abovementioned equation, higher dimensional input dataset is $(x_1, x_2, x_3, \dots, x_m)$, the reconstructed vector of input dataset is $\hat{X}(\hat{x}_1, \hat{x}_2, \hat{x}_3, \dots, \hat{x}_m)$ and lower dimension vector output from the hidden layer $Z = (h_1, h_2, h_3, \dots, h_k)$. Here, g and h indicate the activation function of the hidden and output layer neurons correspondingly. W and W' shows the weight matrix and b and b' indicates bias vector of encoder and decoder correspondingly. The reconstructed error function E between the input x and rebuilt input $\hat{x}$ employs MSE function.

$$E = \frac{1}{N} \sum_{i=1}^N x_i + x_i^2, \quad (4)$$

The smoothed $l1$ regularization was added to ML model. The $l1$ regularization is non-distinguishable and the result is difficult for optimization. Beck and Teboulle developed a novel smoothing method labelled as inf-conv smoothing approach. The inf-conv of the first and second-order overall dissimilarity as the regularization term in penalized supreme probability rennovation.

$$g_u(t) = \begin{cases} \frac{t^2}{2\mu}, & \text{if } |t| \leq \mu \\ |t| - \frac{\mu}{2}, & \text{otherwise} \end{cases} \quad (5)$$

Where $\mu > 0$ denotes the hyperparameter that controls the resemblance between $l1$ and smoothed $l1$ regularization. If $\mu = 0$, then it becomes $l1$ regularization. SAE employs backpropagation to reduce the cost function that is given below

$$J_{\text{sparse}}(W, b) = \left[\frac{1}{N} \sum_{i=1}^N \frac{1}{2} \|h_{W,b}(x_i) - \hat{x}\|^2 \right] + \frac{\lambda}{2} \sum_{l=1}^{n_l-1} \sum_{i=1}^{s_l} \sum_{j=1}^{s_{l+1}} (W_{ji}^{(l)})^2 + \beta \sum_{j=1}^n S(t_j) \quad (6)$$

Where $S(\cdot)$ denotes the function that could be inducing the sparsity.

$$t_j = \frac{1}{m} \sum_{i=1}^m a_j^{(i)} \quad (7)$$

Where, $a_j^{(i)}$ characterizes the output of j^{th} hidden units of i^{th} input samples. By applying the smoothed $l1$ regularization rather than KL divergence, in the abovementioned sparsity function. But, selecting the activation function ReLU.

$$f(x) = \max(0, x) \quad (8)$$

Where $a_j^{(i)}$ the accepted value from 0 to 1. This implies, the average activation t_j of j^{th} hidden unit on the whole trainable

set ranges within $[0, 1]$, the situation of $t > 0$ should be taken into account.

Sparsity constraint is executed in SAE, those are controlling the hidden layers. The sparsity constraint modifies the error condition by adding penalty term for reflecting deviations from desirable sparsity and later employing BP integrating the sparsity penalty. Involvement with dissimilar levels of sparsity indicates the inverse relationships amongst the sparsity level and the relationship taken in the trainable dataset. A high level of sparsity captures very effective features.

C. Parameter Tuning using MGWO Algorithm

To enhance the performance of the SAE model, the MGWO algorithm is employed for hyperparameter tuning process. Mirjalili et al. proposed a GWO which is metaheuristic algorithm used for large number of optimization problems [18]. The mathematical modeling was created and needs the modification of a small number of parameters, which allows exploration and exploitation that attained in a direct way while providing promising convergence.

The motivation for the GWO is predatory behavior of the grey wolf in the wild. They prefer to live in a group with several participants and rigid hierarchy. In this work, the population can be broken down into alpha (α)-group leader, beta (β)-who can take over as leader of the group if the alpha wolf gets sick or dies, and later delta (δ) and omega (ϕ)-the remaining group. Predation includes a sequence of steps: harassment, pursuit, attack, and encirclement behaviors. Every wolf signifies a potential solution, whereas the prey itself is the better option. α , β , and δ indicate the better solution. The residual candidates are noted as ϕ . As per the location of, β , and δ , the position of the omega is upgrading. A wolf (wo) makes use of Eqs. (9) and (12) to determine how farther it is from the three optimum solutions:

$$\vec{A} = 2\vec{a} \cdot \vec{v}_1 - \vec{a} \quad (9)$$

$$\vec{C} = 2\vec{v}_2 \quad (10)$$

$$\vec{D}_\alpha = |\vec{C} \cdot \vec{X}_\alpha - \vec{X}_{wo}|, \vec{D}_\beta = |\vec{C} \cdot \vec{X}_\beta - \vec{X}_{wo}|, \vec{D}_\delta = |\vec{C} \cdot \vec{X}_\delta - \vec{X}_{wo}| \quad (11)$$

$$\vec{X} = \vec{X}_\alpha - \vec{A} \cdot \vec{D}_\alpha \rightarrow \vec{X}_2 = \vec{X}_\beta - \vec{A} \cdot \vec{D}_\beta \rightarrow \vec{X}_3 = \vec{X}_\delta - \vec{A} \cdot \vec{D}_\delta \quad (12)$$

$$\vec{X}_{wo}(nt) \rightarrow \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \quad (13)$$

From the expression, $\vec{X}_{wo}$, $\vec{X}_\alpha$, $\vec{X}_\beta$, and $\vec{X}_\delta$ denotes the location of wolf vectors for wo , α , β , and δ , whereas nt indicates the following iteration of wolf vectors. $\vec{A}$ and $\vec{C}$ represent the coefficient vectors. $\vec{D}_\alpha$, $\vec{D}_\beta$, and $\vec{D}_\delta$ indicates the distance vectors between α , β , δ , and wo . $\vec{v}_1$ and $\vec{v}_2$ represent vectors derived randomly within $[0, 1]$. Therefore, writing an arrow over the parameter is the typical way to write a vector:

$$\vec{a} = \frac{2 - 2 \cdot t}{t_{\max}} \quad (14)$$

In Eq. (6), t indicates the amount of existing iterations, and $t_{\max}$ denotes the maximal amount of iterations.

The MGWO algorithm is derived by the use of quantum-based optimization (QBO) algorithm. Every feature in QBO is characterized as a quantum bit ($Q\text{-bit}(q)$), whereas q indicates the superposition of binary values ('1' and '0'). The

subsequent formula is utilized for establishing the mathematical equation of Q - bit (q).

$$q = \alpha + i\beta = e^{i\theta}, |\alpha|^2 + |\beta|^2 \quad (15)$$

In Eq. (15), the possible value of the Q - bit being 0' and 1' is provided by α and β , correspondingly. The variable θ indicates the angle of q , and it is upgraded by $\tan^{-1}(\alpha/\beta)$.

The procedure of finding the value changing of q is the major objective of QBO, and it is defined as follows by calculating $\Delta\theta$ as:

$$q(t+1) = q(t) \times R(\Delta\theta) = [\alpha(t)\beta(t)] \times R(\Delta\theta) \quad (16)$$

$$R(\Delta\theta) = \begin{Bmatrix} \cos(\Delta\theta) & -\sin(\Delta\theta) \\ \sin(\Delta\theta) & \cos(\Delta\theta) \end{Bmatrix} \quad (17)$$

In Eq. (17), $\Delta\theta$ refers the rotation angle of i^{th} Q - bit of j^{th} Q - solution. The value of $\Delta\theta$ is predetermined as per X_b .

The MGWO manner improves a fitness function (FF) for achieving greater classifier results. It describes a positive integer for demonstrating the good efficiency of candidate outcomes. During this case, the minimizing of the classifier rate of errors is observed as FF is provided in Eq. (18).

$$\text{fitness}(x_i) = \frac{\text{ClassifierErrorRate}(x_i)}{\text{number of misclassified samples}} * 100 \quad (18)$$

IV. RESULTS AND DISCUSSION

Here, the FCP results of the MGWO-SAE method can be examined on two datasets namely Analcatdata and German credit dataset as depicted in Table 1.

TABLE I
DETAILS OF DATASET

Dataset	No. of instances	No. of attributes	No. of class	Bankrupt/Non-Bankrupt
AnalcatData	50	5	2	25/25
German Credit	1000	24	2	300/700

Table 2 reports the FCP outcomes of the MGWO-SAE technique on Analcat dataset. The table values notified that the MGWO-SAE technique has properly recognized the bankrupt and non-bankrupt classes. For instance, on 70% of TR set, the MGWO-SAE technique offers average $accu_y$, $sens_y$, and $spec_y$ of 97.59%, 97.59%, and 97.59% correspondingly. Also, on 30% of TS set, the MGWO-SAE system provides average $accu_y$, $sens_y$, and $spec_y$ of 98.60%, 98.60%, and 98.60% correspondingly.

TABLE II

FCP OUTCOME OF MGWO-SAE APPROACH ON ANALCAT DATASET

Analcat Dataset			
Class Labels	$Accu_y$	$Sens_y$	$Spec_y$
Training Phase (70%)			
Bankrupt	96.64	96.64	98.54
Non-Bankrupt	98.54	98.54	96.64
Average	97.59	97.59	97.59

Testing Phase (30%)			
Bankrupt	97.20	97.20	100.00
Non-Bankrupt	100.00	100.00	97.20
Average	98.60	98.60	98.60

The TACY and VACY of the MGWO-SAE approach on Analcat dataset are represented in Fig. 2. The figure pointed out that the MGWO-SAE system has exposed better performance with higher values of TACY and VACY. It is noticeable that the MGWO-SAE algorithm has gained maximal TACY outcome.

The TLOS and VLOS of the MGWO-SAE algorithm on Analcat dataset are represented in Fig. 3. The figure stated that the MGWO-SAE system has exposed higher performance with minimal values of TLOS and VLOS. It is evident that the MGWO-SAE algorithm has resulted in to decrease in VLOS outcome.

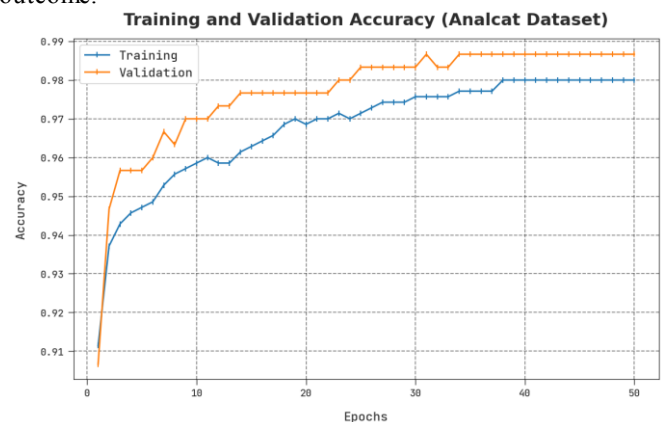


Fig. 2. TACY and VACY outcome of MGWO-SAE approach on Analcat dataset

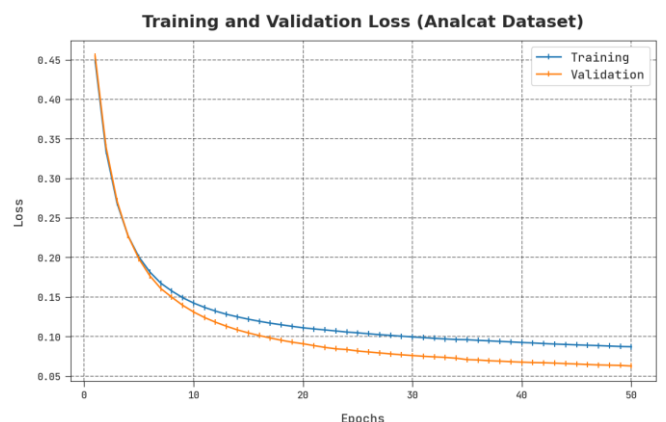


Fig. 3. TLOS and VLOS outcome of MGWO-SAE approach on Analcat dataset

Table 3 defines the FCP outcomes of the MGWO-SAE approach on German Credit Dataset. The table values stated that the MGWO-SAE technique has properly recognized the bankrupt and non-bankrupt classes. For sample, on 70% of TR set, the MGWO-SAE algorithm provides average $accu_y$, $sens_y$, and $spec_y$ of 98.15%, 98.15%, and 98.15% correspondingly. Simultaneously, on 30% of TS set, the

MGWO-SAE method suggests average $accu_y$, $sens_y$, and $spec_y$ of 98.95%, 98.95%, and 98.95% correspondingly.

TABLE III

FCP OUTCOME OF MGWO-SAE APPROACH ON GERMAN CREDIT DATASET

German Credit Dataset			
Class Labels	$Accu_y$	$Sens_y$	$Spec_y$
Training Phase (70%)			
Bankrupt	97.76	97.76	98.54
Non-Bankrupt	98.54	98.54	97.76
Average	98.15	98.15	98.15
Testing Phase (30%)			
Bankrupt	97.90	97.90	100.00
Non-Bankrupt	100.00	100.00	97.90
Average	98.95	98.95	98.95

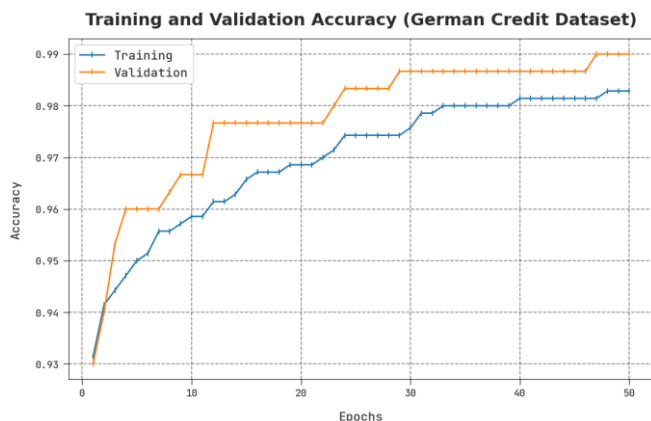


Fig. 4. TACY and VACY outcome of MGWO-SAE approach on German Credit dataset

The TACY and VACY of the MGWO-SAE approach on German Credit dataset are signified in Fig. 4. The figure stated that the MGWO-SAE methodology has depicted enhanced performance with enhanced values of TACY and VACY. It is noticeable that the MGWO-SAE algorithm has reached maximal TACY outcome.

The TLOS and VLOS of the MGWO-SAE system on German Credit dataset are portrayed in Fig. 5. The figure stated that the MGWO-SAE approach has depicted superior performance with decreased values of TLOS and VLOS. It is observable that the MGWO-SAE algorithm has resulted in minimal VLOS outcome.

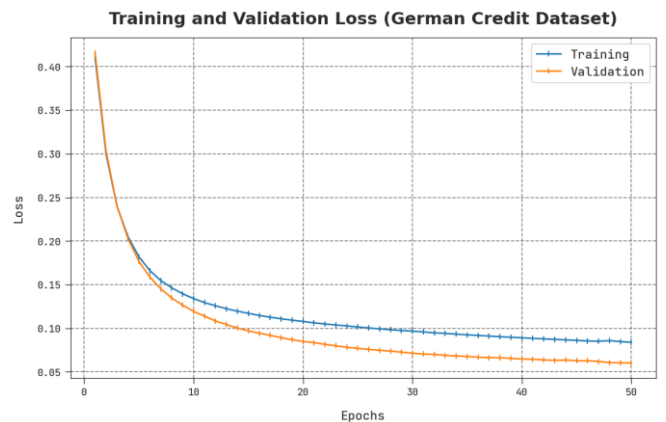


Fig. 5. TLOS and VLOS outcome of MGWO-SAE approach on German Credit dataset

Table 4 and Fig. 6 inspect the results of the MGWO-SAE technique with other FCP approaches on Analcat dataset [15]. The results assured the supremacy of the MGWO-SAE technique. Based on $accu_y$, the MGWO-SAE technique results in improved $accu_y$ of 98.60% while the MLP, SVM, ACO, AdaBoost, and LSTM-RNN models accomplish decreased $accu_y$ of 96.85%, 95.76%, 97.39%, 94.93%, and 94.21% respectively. Meanwhile, with respect to $sens_y$, the MGWO-SAE system results in improved $sens_y$ of 98.60% while the MLP, SVM, ACO, AdaBoost, and LSTM-RNN approaches realize reduced $sens_y$ of 97.69%, 95.47%, 97.90%, 94.49% and 95.91% correspondingly. Eventually, in terms of $spec_y$, the MGWO-SAE method results in higher $spec_y$ of 98.60% while the MLP, SVM, ACO, AdaBoost, and LSTM-RNN methodologies accomplish lesser $spec_y$ of 97.46%, 97.57%, 97.79%, 95.62% and 96.80% correspondingly.

TABLE IV

COMPARATIVE OUTCOME OF MGWO-SAE APPROACH WITH OTHER SYSTEMS UNDER ANALCAT DATASET

Analcat Dataset			
Methods	Accuracy	Sensitivity	Specificity
MGWO-SAE	98.60	98.60	98.60
MLP	96.85	97.69	97.46
SVM	95.76	95.47	97.57
ACO	97.39	97.90	97.79
AdaBoost	94.93	94.49	95.62
LSTM-RNN	94.21	95.91	96.80

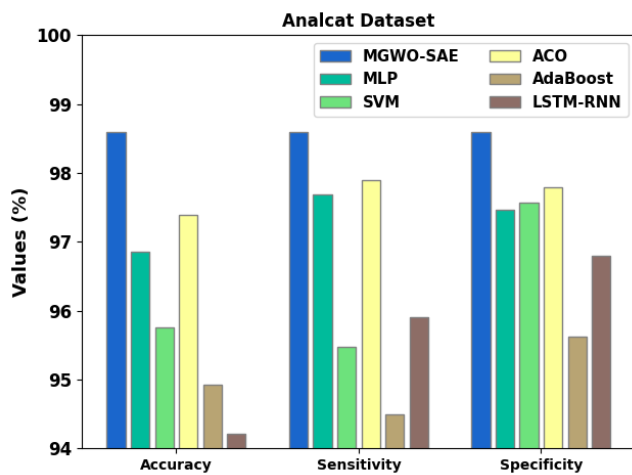


Fig. 6. Comparative analysis of MGWO-SAE approach under Analcat dataset

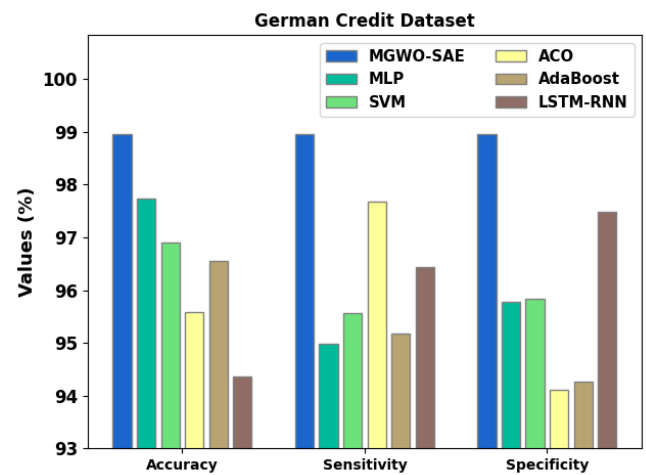


Fig. 7. Comparative analysis of MGWO-SAE approach under German Credit dataset

TABLE V

COMPARATIVE OUTCOME OF MGWO-SAE ALGORITHM WITH OTHER SYSTEMS UNDER GERMAN CREDIT DATASET

German Credit Dataset			
Methods	Accuracy	Sensitivity	Specificity
MGWO-SAE	98.95	98.95	98.95
MLP	97.74	94.98	95.78
SVM	96.91	95.57	95.83
ACO	95.59	97.67	94.11
AdaBoost	96.56	95.17	94.27
LSTM-RNN	94.37	96.43	97.48

Table 5 and Fig. 7 examine the outcome of the MGWO-SAE method with other FCP approaches on German Credit dataset. The outcomes referred to the supremacy of the MGWO-SAE technique. With respect to $accu_y$, the MGWO-SAE technique results in improved $accu_y$ of 98.95% while the MLP, SVM, ACO, AdaBoost, and LSTM-RNN models accomplish decreased $accu_y$ of 97.74%, 96.91%, 95.59%, 96.56% and 94.37% respectively. Meanwhile, with respect to $sens_y$, the MGWO-SAE method results in superior $sens_y$ of 98.95% while the MLP, SVM, ACO, AdaBoost, and LSTM-RNN systems realize decreased $sens_y$ of 94.98%, 95.57%, 97.67%, 95.17% and 96.43% correspondingly. Finally, in terms of $spec_y$, the MGWO-SAE technique results in enhanced $spec_y$ of 98.95% while the MLP, SVM, ACO, AdaBoost, and LSTM-RNN systems achieve minimal $spec_y$ 95.78%, 95.83%, 94.11%, 94.27% and 97.48% correspondingly.

These results confirmed the enhanced performance of the MGWO-SAE technique over other existing models.

V. CONCLUSION

In this article, we have developed a novel MGWO-SAE approach for automated FCP in small marginal firms. The goal of the MGWO-SAE technique lies in the effectual forecasting of financial crisis in the small marginal firm. To achieve this, the presented MGWO-SAE technique uses data preprocessing to transform the input financial data into a proper format. For prediction purposes, the MGWO-SAE technique uses SAE classification method. To enhance the performance of the SAE model, the MGWO algorithm is employed for hyperparameter tuning model. The experimental result analysis of the MGWO-SAE technique is tested on financial dataset. The major limitations are feature prediction. The experimental outcomes highlighted the improvements of the MGWO-SAE technique over other models. In future, feature reduction strategy can be developed to improve the predictive results.

REFERENCES

- [1] Zeng, H., 2022. Influences of mobile edge computing-based service preloading on the early-warning of financial risks. *The Journal of Supercomputing*, 78(9), pp.11621-11639.
- [2] Schmitz, T., 2021. Endogenous growth, firm heterogeneity and the long-run impact of financial crises. *European Economic Review*, 132, p.103637.
- [3] Kobina Enos, B., Yensu, J. and Obeng, H., 2020. Global financial crisis and determinants of capital structure: Evidence from Ghanaian non-financial listed firms. *The International Journal of Business and Finance Research*, 14(1), pp.35-56.
- [4] Tran, K.L., Le, H.A., Nguyen, T.H. and Nguyen, D.T., 2022. Explainable Machine Learning for Financial Distress Prediction: Evidence from Vietnam. *Data*, 7(11), p.160.
- [5] Bianchi, J. and Mendoza, E.G., 2020. A fisherian approach to financial crises: Lessons from the sudden stops literature. *Review of Economic Dynamics*, 37, pp.S254-S283.
- [6] Buckmann, M., Gallego Marquez, P., Gimpelewicz, M., Kapadia, S. and Rismanchi, K., 2021. The more the merrier? Evidence from the global financial crisis on the value of multiple requirements in bank regulation.
- [7] Hasumi, R., Iiboshi, H., Matsumae, T. and Nakamura, D., 2019. Does a financial accelerator improve forecasts during Financial crises? Evidence from Japan with prediction-pooling methods. *Journal of Asian Economics*, 60, pp.45-68.

- [8] Chen, J., Ma, S. and Wu, Y., 2022. International carbon financial market prediction using particle swarm optimization and support vector machine. *Journal of Ambient Intelligence and Humanized Computing*, 13(12), pp.5699-5713.
- [9] Du, Q., 2021. FinancialCrisis Early Warning Based on Panel Data and Dynamic Dual Choice Model. *Complexity*, 2021.
- [10] Bocola, L. and Lorenzoni, G., 2020. Financial crises, dollarization, and lending of last resort in open economies. *American Economic Review*, 110(8), pp.2524-57.
- [11] Zhang, C., Zhong, H. and Hu, A., 2022. Research on Early Warning of Financial Crisis of Listed Companies Based on Random Forest and Time Series. *Mobile Information Systems*, 2022.
- [12] Chen, Y.J. and Wu, C.Y., 2021. Predicting a corporate financial crisis using letters to shareholders. *Soft Computing*, 25(5), pp.3623-3636.
- [13] Batuman, B., Yildiz, Y. and Karan, M.B., 2022. The impact of the global financial crisis on corporate cash holdings: Evidence from Eastern European countries. *Borsa Istanbul Review*, 22(4), pp.678-687.
- [14] Dastkhan, H., 2021. Network- based early warning system to predict financial crisis. *International Journal of Finance & Economics*, 26(1), pp.594-616.
- [15] Tyagi, S.K.S. and Boyang, Q., 2021. An intelligent internet of things aided financial crisis prediction model in fintech. *IEEE Internet of Things Journal*.
- [16] Lu, X., He, W., Lu, Q. and Tang, Y., 2021, November. Hybrid Filter-wrapper Feature Selection Using Water Wave Optimization for Financial Crisis Prediction in Enterprises. In *2021 16th International Conference on Intelligent Systems and Knowledge Engineering (ISKE)* (pp. 193-199). IEEE.
- [17] Rao, K.N., Rao, K.V. and PVGD, P.R., 2021. A hybrid intrusion detection system based on sparse autoencoder and deep neural network. *Computer Communications*, 180, pp.77-88.
- [18] Elaziz, M.A., Ewees, A.A., Al-qaness, M.A., Alshathri, S. and Ibrahim, R.A., 2022. Feature Selection for High Dimensional Datasets Based on Quantum-Based Dwarf Mongoose Optimization. *Mathematics*, 10(23), p.4565.